

Frequently Asked Questions

2016 Workplace and Gender Relations Survey of Active Duty Members Office of People Analytics

The Defense Research, Surveys, and Statistics Center (RSSC), Office of People Analytics (OPA), has been conducting surveys of gender issues for the active duty military since 1988. RSSC uses scientific state of the art statistical techniques to draw conclusions from random, representative samples of the active duty populations. To construct estimates for the 2016 Workplace and Gender Relations Survey of Active Duty Members (2016 WGRA), OPA used complex sampling and weighting procedures to ensure accuracy of estimates to the full active duty population. This approach, though widely accepted as the standard method to construct generalizable estimates, is often misunderstood. The following details some common questions about our methodology as a whole and the 2016 WGRA specifically.

1. What was the population of interest for the 2016 WGRA?

The target population consisted of members from the active duty from the Army, Navy, Marine Corps, Air Force, and Coast Guard who were below flag rank and have at least four months of service.

OPA sampled 50% of men and 75% of women, consisting of 735,329 members. Data were collected between 22 July and 14 October 2016.

The weighted total DoD response rate for the 2016 WGRA was 23%, which is typical for large DoD-wide surveys. This rate was similar to the 29% response rate for the 2014 Rand Military Workplace Survey and the 24% response rate in the 2012 Workplace and Gender Relations Survey of Active Duty Members.

2. The 2016 WGRA uses “sampling” and “weighting.” Why are these methods used and what do they do?

Simply stated, sampling and weighting allows for data, based on a sample, to be accurately generalized up to the total population. In the case of the 2016 WGRA, this allows OPA to generalize to the full population of active duty members that meet the criteria listed above.

In stratified random sampling, all members of a population are categorized into homogeneous groups. For example, members might be grouped by gender and component (e.g., all male Army personnel in one group, all female Army personnel in another). Members are chosen at random within each group so that all eligible military members have an equal chance of selection to participate in the survey. Small groups are oversampled in comparison to their proportion of the population so there will be enough responses (approximately 500) from small groups to provide reliable estimates for population subgroups.

OPA scientifically weights the data so findings can be generalized to the full population of active duty members. Within this process, statistical adjustments are made to ensure the sample more accurately reflects the characteristics of the population from which it was drawn. This ensures that the oversampling within any one subgroup does not result in overrepresentation in the total force estimates, and also properly adjusts to account for survey nonresponse.

This methodology meets industry standards used by government statistical agencies including the Census Bureau, Bureau of Labor Statistics, National Agricultural Statistical Service, National Center for Health Statistics, and National Center for Education Statistics. In addition, private survey firms including RAND, WESTAT, and RTI use this methodology, as do well-known polling firms such as Gallup, Pew, and Roper.

3. **Are survey estimates valid with only a 23% weighted response rate?**

Response rates to the 2016 WGRA are consistent with response rate levels and trends for both the 2014 Rand Military Workplace Survey (29% response rate) and the 2012 Workplace and Gender Relations Survey of Active Duty Members (24% response rate). Experts in the field have found that surveys with similar response rates, or lower, are able to produce reliable estimates. While non-response bias due to low response rates is always a concern, OPA has knowledge, based on administrative records, of the characteristics of both survey respondents and survey nonrespondents, and uses this information to make statistical adjustments that compensate for survey non-response. This important advantage improves the quality of estimates from OPA surveys that other survey organizations rarely have.

OPA uses accurate administrative records (e.g., demographic data) for the active duty population both at the sample design stage as well as during the statistical weighting process to account for survey non-response and post-stratification to known key variables or characteristics. Prior OPA surveys provide empirical results showing how response rates vary by many characteristics (e.g., paygrade and Service). OPA uses this information to accurately estimate the optimum sample sizes needed to obtain sufficient numbers of respondents within key reporting groups (e.g., Army, female). After the survey is complete, OPA makes statistical weighting adjustments so that each subgroup (e.g., Army, E1-E3, and female) contributes toward the survey estimates proportional to the known size of the subgroup.

In addition, OPA routinely conducts “Non-Response Bias Analyses” on the Gender Relations surveys. This type of analyses measures whether respondents to the survey are fundamentally different from non-responders on a variety of dimensions. If differences are found, this may be an indication that there is bias in the estimates produced. Using a variety of methods to gauge potential non-response bias, OPA has found no evidence of non-response bias on the Gender Relations Surveys (OPA, 2016a).

4. **Is 23% a common response rate for other military or civilian surveys?**

Response rates of less than 30% are not uncommon for surveys that use similar sampling and weighting procedures. Many civilian surveys often do not have the same knowledge about the composition of the total population in order to generalize results to the full population via sampling and weighting. Therefore, these surveys often require much higher response rates in order to construct accurate estimates. For this reason, it is difficult to compare civilian survey response rates to OPA survey response rates. However, many of the large-scale surveys conducted by DoD or civilian survey agencies rely on similar sampling and weighting procedures as OPA to obtain accurate and generalizable findings with response rates lower than 30% (see Q5). Of note, OPA has a further advantage over these surveys by maintaining the administrative record data (e.g., demographic data) on the full population. This rich data, rarely available to survey organizations, is used to reduce bias associated with the weighted estimates and increase the precision and accuracy of estimates.

5. **Can you give some examples of other studies with similar response rates that were used by DoD to understand military populations and inform policy?**

The 2011 Health and Related Behaviors Survey, conducted by ICF International on behalf of the Tricare Activity Management, had a 22% response rate weighted up to the full active duty military population. This 22% represented approximately 34,000 respondents from a sample of about 154,000 active duty military members. In 2010, Gallup conducted a survey for the Air Force on sexual assault within the Service. Gallup weighted the results to generalize to the full population of Air Force members based on about 19,000 respondents representing a 19% response rate. Finally, in 2011, the U.S. Department of Defense Comprehensive Review Working Group, with the assistance of Westat and OPA, conducted a large-scale survey to measure the impact of overturning the Don't Ask Don't Tell (DADT) policy. The DADT survey, which was used to inform DoD policy, was sent to 400,000 active duty and Reserve members. It had a 28% response rate and was generalized up to the full population of military members, both active duty and Reserve. The survey methodology used for this survey, which used the OPA sampling design, won the 2011 Policy Impact Award from The American Association for Public Opinion Research (AAPOR), which “recognizes outstanding research that has had a clear impact on improving policy decisions practice or discourse, either in the public or private sectors.”

6. **What about surveys that study the total U.S. population? How do they compare?**

Surveys of sensitive topics and rare events rely on similar methodology and response rates to project estimates to the total U.S. adult population. For example, the 2010 National Intimate Partner and Sexual Violence Survey, conducted by the Centers for Disease Control and Prevention, calculated population estimates on a variety of sensitive measures based on about 18,000 interviews, reflecting a weighted response rate of between 28% to 34%.

